

J. Grun

#8

1641

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/125,031DATE: 11/15/1999
TIME: 14:00:30

Input Set: I125031.RAW

This Raw Listing contains the General Information
Section and up to first 5 pages.

ENTERED

```
1  <110> APPLICANT: LONGACRE-ANDRE, SHIRLEY
2      ROTH, CHARLES
3      NATO, FARIDABANO
4      BARNWELL, JOHN
5      MENDIS, KAMINI
6  <120> TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF
7      PLASMODIUM MSP-1
8  <130> FILE REFERENCE: 0660-0139-0XPCT
9  <140> CURRENT APPLICATION NUMBER: US/09/125,031
10 <141> CURRENT FILING DATE: 1999-03-10
11 <150> EARLIER APPLICATION NUMBER: PCT/FR97/00290
12 <151> EARLIER FILING DATE: 1997-02-14
13 <150> EARLIER APPLICATION NUMBER: FR96/01822
14 <151> EARLIER FILING DATE: 1996-02-14
15 <160> NUMBER OF SEQ ID NOS: 14
16 <170> SOFTWARE: PatentIn Ver. 2.1
17 <210> SEQ ID NO 1
18 <211> LENGTH: 291
19 <212> TYPE: DNA
20 <213> ORGANISM: Artificial Sequence
21 <220> FEATURE:
22 <223> OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
23 <220> FEATURE:
24 <221> NAME/KEY: CDS
25 <222> LOCATION: (1)..(291)
26 <400> SEQUENCE: 1
27      gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag      48
28      Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
29      1              5              10              15
30      aac tct ggc tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt      96
31      Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
32      20              25              30
33      ctg ctg aac tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac      144
34      Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
35      35              40              45
36      ccg acc tgt aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc      192
37      Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
38      50              55              60
39      acc gag gag gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt      240
40      Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
41      65              70              75              80
42      acc aaa ccc gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc taa      288
43      Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
44      85              90              95
```

PAGE: 2

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/125,031

DATE: 11/15/1999
TIME: 14:00:30

Input Set: I125031.RAW

```

45      taa                                                    291
46  <210> SEQ ID NO 2
47  <211> LENGTH: 95
48  <212> TYPE: PRT
49  <213> ORGANISM: Artificial Sequence
50  <220> FEATURE:
51  <223> OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
52  <400> SEQUENCE: 2
53      Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
54      1              5              10              15
55      Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
56      20              25              30
57      Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
58      35              40              45
59      Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
60      50              55              60
61      Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
62      65              70              75              80
63      Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
64      85              90              95
65  <210> SEQ ID NO 3
66  <211> LENGTH: 279
67  <212> TYPE: DNA
68  <213> ORGANISM: Plasmodium falciparum
69  <400> SEQUENCE: 3
70      aacatttcac aacaccaatg cgtaaaaaaa caatgtccag aaaattcttg atgtttcaga 60
71      catttagatg aaagagaaga atgtaaatgt ttattaaatt acaaacaaga aggtgataaa 120
72      tgtgttgaaa atccaaatcc tacttgtaac gaaaataatg gtggatgtga tgcagatgcc 180
73      aaatgtaccg aagaagattc aggtagcaac ggaaagaaaa tcacatgtga atgtactaaa 240
74      cctgattctt atccactttt cgatggtatt ttctgcagt                279
75  <210> SEQ ID NO 4
76  <211> LENGTH: 354
77  <212> TYPE: DNA
78  <213> ORGANISM: Artificial Sequence
79  <220> FEATURE:
80  <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC
81  <220> FEATURE:
82  <221> NAME/KEY: CDS
83  <222> LOCATION: (1)..(354)
84  <400> SEQUENCE: 4
85      gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag      48
86      Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
87      1              5              10              15
88      aac tct ggc tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt      96
89      Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
90      20              25              30
91      ctg ctg aac tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac      144
92      Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
93      35              40              45
94      ccg acc tgt aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc      192

```

PAGE: 3

RAW SEQUENCE LISTING PATENT APPLICATION US/09/125,031

DATE: 11/15/1999
TIME: 14:00:30

Input Set: I125031.RAW

```

95      Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
96          50                      55                      60
97      acc gag gag gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt      240
98      Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
99          65                      70                      75                      80
100     acc aaa ccc gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc tcc      288
101     Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser Ser
102          85                      90                      95
103     tct aac ttc ttg ggc atc tcg ttc ttg ttg atc ctc atg ttg atc ttg      336
104     Ser Asn Phe Leu Gly Ile Ser Phe Leu Leu Ile Leu Met Leu Ile Leu
105          100                      105                      110
106     tac agc ttc att taa taa
107     Tyr Ser Phe Ile
108          115
109     <210> SEQ ID NO 5
110     <211> LENGTH: 116
111     <212> TYPE: PRT
112     <213> ORGANISM: Artificial Sequence
113     <220> FEATURE:
114     <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC
115     <400> SEQUENCE: 5
116     Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
117         1                      5                      10                      15
118     Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
119         20                      25                      30
120     Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
121         35                      40                      45
122     Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
123         50                      55                      60
124     Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
125         65                      70                      75                      80
126     Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser Ser
127         85                      90                      95
128     Ser Asn Phe Leu Gly Ile Ser Phe Leu Leu Ile Leu Met Leu Ile Leu
129         100                      105                      110
130     Tyr Ser Phe Ile
131         115
132     <210> SEQ ID NO 6
133     <211> LENGTH: 342
134     <212> TYPE: DNA
135     <213> ORGANISM: Plasmodium falciparum
136     <400> SEQUENCE: 6
137     aacattttcac aacaccaatg cgtaaaaaaa caatgtccag aaaattcttg atgtttcaga 60
138     catttagatg aaagagaaga atgtaaatgt ttattaaatt acaaacaaga aggtgataaa 120
139     tgtgttgaaa atccaaatcc tacttgtaac gaaaataatg gtggatgtga tgcagatgcc 180
140     aaatgtaccg aagaagattc aggtagcaac ggaaagaaaa tcacatgtga atgtactaaa 240
141     cctgattctt atccactttt cgatggattt ttctgcagtt cctctaactt cttaggaata 300
142     tcattcttat taatactcat gttaatatta tacagtttca tt                        342
143     <210> SEQ ID NO 7
144     <211> LENGTH: 387

```

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/125,031

DATE: 11/15/1999
TIME: 14:00:30

Input Set: I125031.RAW

```

145 <212> TYPE: DNA
146 <213> ORGANISM: Plasmodium falciparum
147 <220> FEATURE:
148 <221> NAME/KEY: CDS
149 <222> LOCATION: (1)..(387)
150 <400> SEQUENCE: 7
151   atg aag gcg cta ctc ttt ttg ttc tct ttc att ttt ttc gtt acc aaa   48
152   Met Lys Ala Leu Leu Phe Leu Phe Ser Phe Ile Phe Phe Val Thr Lys
153       1               5               10               15
154   gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag   96
155   Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
156       20               25               30
157   gaa ttc aac atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag   144
158   Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
159       35               40               45
160   aac tct ggc tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt   192
161   Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
162       50               55               60
163   ctg ctg aac tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac   240
164   Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
165       65               70               75               80
166   ccg acc tgt aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc   288
167   Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
168       85               90               95
169   acc gag gag gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt   336
170   Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
171       100              105              110
172   acc aaa ccc gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc taa   384
173   Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
174       115              120              125
175   taa   387
176 <210> SEQ ID NO 8
177 <211> LENGTH: 127
178 <212> TYPE: PRT
179 <213> ORGANISM: Plasmodium falciparum
180 <400> SEQUENCE: 8
181   Met Lys Ala Leu Leu Phe Leu Phe Ser Phe Ile Phe Phe Val Thr Lys
182       1               5               10               15
183   Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
184       20               25               30
185   Glu Phe Asn Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu
186       35               40               45
187   Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys
188       50               55               60
189   Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn
190       65               70               75               80
191   Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys
192       85               90               95
193   Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys
194       100              105              110

```

PAGE: 5

RAW SEQUENCE LISTING PATENT APPLICATION US/09/125,031

DATE: 11/15/1999
TIME: 14:00:30

Input Set: I125031.RAW

```

195      Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
196              115              120              125
197 <210> SEQ ID NO 9
198 <211> LENGTH: 330
199 <212> TYPE: DNA
200 <213> ORGANISM: Plasmodium falciparum
201 <220> FEATURE:
202 <221> NAME/KEY: CDS
203 <222> LOCATION: (1)..(330)
204 <400> SEQUENCE: 9
205      gaa aca gaa agt tat aag cag ctt gta gcc aac gtg gac gaa ttc aac      48
206      Glu Thr Glu Ser Tyr Lys Gln Leu Val Ala Asn Val Asp Glu Phe Asn
207              1              5              10              15
208      atc tcg cag cac caa tgc gtg aaa aaa caa tgt ccc gag aac tct ggc      96
209      Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly
210              20              25              30
211      tgt ttc aga cac ttg gac gag aga gag gag tgt aaa tgt ctg ctg aac      144
212      Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn
213              35              40              45
214      tac aaa cag gag ggc gac aag tgc gtg gag aac ccc aac ccg acc tgt      192
215      Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn Pro Thr Cys
216              50              55              60
217      aac gag aac aac ggc ggc tgt gac gca gac gcc aaa tgc acc gag gag      240
218      Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys Thr Glu Glu
219              65              70              75              80
220      gac tcg ggc agc aac ggc aag aaa atc acg tgt gag tgt acc aaa ccc      288
221      Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys Thr Lys Pro
222              85              90              95
223      gac tcg tac ccg ctg ttc gac ggc atc ttc tgc agc taa taa      330
224      Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
225              100              105              110
W--OK
226 <210> SEQ ID NO 10
227 <211> LENGTH: 108
228 <212> TYPE: PRT
229 <213> ORGANISM: Plasmodium falciparum
230 <400> SEQUENCE: 10
231      Glu Thr Glu Ser Tyr Lys Gln Leu Val Ala Asn Val Asp Glu Phe Asn
232              1              5              10              15
233      Ile Ser Gln His Gln Cys Val Lys Lys Gln Cys Pro Glu Asn Ser Gly
234              20              25              30
235      Cys Phe Arg His Leu Asp Glu Arg Glu Glu Cys Lys Cys Leu Leu Asn
236              35              40              45
237      Tyr Lys Gln Glu Gly Asp Lys Cys Val Glu Asn Pro Asn Pro Thr Cys
238              50              55              60
239      Asn Glu Asn Asn Gly Gly Cys Asp Ala Asp Ala Lys Cys Thr Glu Glu
240              65              70              75              80
241      Asp Ser Gly Ser Asn Gly Lys Lys Ile Thr Cys Glu Cys Thr Lys Pro
242              85              90              95
243      Asp Ser Tyr Pro Leu Phe Asp Gly Ile Phe Cys Ser
244              100              105

```

PAGE: 6

VERIFICATION SUMMARY
PATENT APPLICATION US/09/125,031

DATE: 11/15/1999
TIME: 14:00:30

Input Set: I125031.RAW

Line ? Error/Warning

Original Text

225 W Invalid/Missing Amino Acid Numbering

100

105